AMENDMENTS

Amendments to the Specification

Please replace paragraph 55 beginning on page 15, line 23 and ending on page 16, line 1 with the following amended paragraph:

In another embodiment, the bound DNA-protein complexes are treated with sonication. Sonication uses high-frequency sound waves to break the non-bound portions of the DNA strands. L. Stryer, *Biochemistry*, 4th Ed., W.H. Freeman and Co., New York, (March 1995): 271. Standard protocols for sonication are described in Chapter 12, Protocol 1 of *Molecular Cloning: A Laboratory Manual* (3rd ed.), Sambrook et al., Vols. 1-3, Cold Spring Harbor Laboratory Press, NY, (2001). Additional protocols for the sonication of DNA specifically are described by Richard Young at *Genome-wide Location and Function of DNA Binding Proteins*. Richard Young. (2000). Massachusettes Institute of Technology. June 25, 2001 ~~<http://web.wi.mit.edu/young/location>~~.

Please replace paragraphs 66 and 67 on page 19 with the following amended paragraphs:

~~Microarray~~ Microarrays can be used in a variety of ways. A preferred microarray contains nucleic acids and is used to analyze nucleic acid samples. Typically, a nucleic acid sample is prepared from an appropriate source and labeled with a signal moiety, such as a fluorescent label. The sample is hybridized with the array under appropriate conditions. The arrays are washed or otherwise processed to remove non-hybridized sample nucleic acids. The hybridization is then evaluated by detecting the distribution of the label on the chip. The distribution of label may be detected by scanning the arrays to

determine fluorescence intensity distribution. Typically, the hybridization of each probe is reflected by several pixel intensities. The raw intensity data may be stored in a gray scale pixel intensity file. The GATC™ Consortium has specified several file formats for storing array intensity data. The final software specification is available at the GATC Consortium's website ~~www.gatcconsortium.org~~ and is incorporated herein by reference in its entirety. The pixel intensity files are usually large. For example, a GATC™ compatible image file may be approximately 50 Mb if there are about 5000 pixels on each of the horizontal and vertical axes and if a two byte integer is used for every pixel intensity. The pixels may be grouped into cells. (See GATC™ software specification). The probes in a cell are designed to have the same sequence; i.e., each cell is a probe area. A CEL file contains the statistics of a cell, e.g., the 75th percentile and standard deviation of intensities of pixels in a cell. The 50, 60, 70, 75 or 80th percentile of pixel intensity of a cell is often used as the intensity of the cell.

The Affymetrix® Analysis Data Model (AADM) is the relational database schema Affymetrix uses to store experiment results. It includes tables to support mapping, spotted arrays and expression results. Affymetrix publishes AADM to support open access to experiment information generated and managed by Affymetrix® software so that results may be filtered and mined with any compatible analysis tools. The AADM specification (Affymetrix, Santa Clara, CA, 2001) is incorporated herein by reference for all purposes. ~~The specification is available at http://www.affymetrix.com/support/aadm/aadm.html, last visited on 9/4/2001.~~